Title of the Invention

AVIAN TRANSGENESIS USING A CHICKEN OVALBUMIN GENE REGION

The present application claims priority from U.S. provisional patent applications, Serial Nos. 60/462,953, filed April 15, 2003; 60/465,015, filed April 24, 2003; and 60/469,488 filed May 9, 2003, all of which are hereby incorporated by reference herein in their entireties.

Field of the Invention

The present invention relates generally to an isolated nucleic acid molecule comprising an avian ovalbumin transcriptional regulatory control region and linked matrix attachment regions. The invention further relates to recombinant nucleic acids and expression vectors, genetically transformed cells and transgenic avians that comprise an avian ovalbumin transcriptional regulatory region operably linked to a heterologous polypeptide-encoding nucleic acid insert. The present invention also relates to the expression and production of the polypeptide-encoding nucleic acid molecule under the control of the isolated avian ovalbumin transcriptional regulatory region.

Background

Transgenic technology to convert animals into "bioreactor" for the production of specific proteins or other substances of pharmaceutical interest (Gordon *et al.*, 1987, *Biotechnology* 5: 1183-1187; Wilmut *et al.*, 1990, *Theriogenology* 33: 113-123) offers significant advantages over more conventional methods of protein production by gene expression.

Recombinant nucleic acid molecules have been engineered so that an expressed heterologous protein may be joined to a protein or peptide that allows secretion of the transgenic expression product into milk or urine, from which the protein may then be recovered. These procedures may require lactating animals, with the attendant costs of maintaining individual animals or herds of large species, such as cows, sheep, or goats.

Historically, transgenic animals have been produced almost exclusively by microinjection of the fertilized egg. The pronuclei of fertilized eggs are microinjected *in vitro* with foreign, i.e., xenogeneic or allogeneic, heterologous DNA or hybrid DNA molecules. The microinjected fertilized eggs are then transferred to the genital tract of a pseudopregnant female (e.g., Krimpenfort *et al.*, U.S. Pat. No. 5,175,384).

One system that holds potential is the avian reproductive system. The production of an avian egg begins with formation of a large yolk in the ovary of the hen. The unfertilized oocyte or ovum is positioned on top of the yolk sac. After ovulation, the ovum passes into the infundibulum of the oviduct where it is fertilized if sperm are present, and then moves into the magnum of the oviduct, which is lined with tubular gland cells. These cells secrete the egg-white proteins, including ovalbumin, lysozyme, ovomucoid, conalbumin and ovomucin, into the lumen of the magnum where they are deposited onto the avian embryo and yolk.

The hen oviduct offers outstanding potential as a protein bioreactor because of the high levels of protein production, the promise of proper folding and post-translation modification of the target protein, the ease of product recovery, and the shorter developmental period of chickens compared to other potential animal species. The chicken ovalbumin gene is highly expressed in the tubular glands of the mature hen oviduct and is therefore a suitable candidate for an efficient promoter for heterologous protein production in transgenic birds. Efforts have been made to create transgenic chickens expressing heterologous proteins in the oviduct by means of microinjection of DNA (PCT Publication WO 97/47739).

Gene expression must be considered not only from the perspective of cis-regulatory elements associated with a gene, and their interactions with trans-acting elements, but also with regard to the genetic environment in which they are located. Chromosomal positioning effects result in variations in levels of transgene expression associated with different locations of the transgene within the recipient genome. An important factor governing the level of transgene expression is the chromatin structure around a transgene, and how it cooperates with the cis-regulatory elements. While the deletion of a cis-regulatory element from a transgenic lysozyme locus can

be sufficient to reduce or eliminate positional independence of the level of gene expression , there is also evidence that positional independence conferred on a transgene requires the cotransfer of many kilobases of DNA other than just the protein encoding region and the immediate cis-transcriptional regulatory elements.

Scattered throughout the chicken genome, including the chicken ovalbumin locus, are short sequences that resemble features of Long Terminal Repeats (LTRs) of retrovirus. The function of these elements is unclear but most likely may help define the DNAse hypersensitive (DHS) regions of a gene locus (Stein *et al.*, 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80: 6485-6489). Thus, flanking various avian genes are matrix attachment regions (5' and 3' MARs), alternatively referred to as "scaffold attachment regions" or SARs. The outer boundaries of the chicken lysozyme locus, for example, have been defined by the MARs (Phi-Van *et al.*, 1988, *E.M.B.O.J.* 7: 655-664; Phi-Van & Stratling., 1996, *Biochem.* 35: 10735-10742). Deletion of a 1.32 kb or a 1.45 kb region, each comprising half of a 5' MAR, reduces positional variation in the level of transgene expression (Phi-Van & Stratling, *supra*).

The 5' matrix attachment region (5' MAR), located about −11.7 kb upstream of the chicken lysozyme transcription start site, can increase the level of gene expression by limiting the chromosomal positional effects exerted against a transgene (Phi-Van *et al.*, 1988, *supra*). At least one other MAR is located 3' downstream of the protein encoding region. Although MAR nucleic acid sequences are conserved, little cross-hybridization is seen, indicating significant overall sequence variation. However, MARs of different species can interact with the nucleomatrices of heterologous species, to the extent, for example, that the chicken lysozyme MAR can associate with the plant tobacco nucleomatrix as well as that of the chicken oviduct cells (Mlynarona *et al.*, 1994, *Cell* 6: 417-426; von Kries *et al.*, 1990, *Nucleic Acids Res.* 18: 3881-3885). The lysozyme promoter region of chicken is also active when transfected into mouse fibroblast cells and linked to a reporter gene such as the bacterial chloramphenicol acetyltransferase gene. In each case, the presence of a 5' MAR element increased positional independency of the level of transcription (Stief *et al.*, 1989, *Nature* 341: 343-345; Sippel *et al.*, pgs. 257 – 265 in Houdeline L.M. (ed), "Transgenic Animals: Generation and Use").

The ability to direct the insertion of a transgene into a site in the genome of an animal where the positional effect is limited offers predictability of results during the development of a desired transgenic animal, and increased yields of the expressed product. Sippel and Steif disclose, in U.S. Patent No. 5,731,178, methods to increase the expression of genes introduced into eukaryotic cells by flanking a transcription unit with scaffold attachment elements, in particular the 5' MAR isolated from the chicken lysozyme gene. The transcription unit disclosed by Sippel and Steif was an artificial construct that combined only the -6.1 kb enhancer element and the proximal promoter element (base position -579 to +15) from the lysozyme gene. Other promoter associated elements were not included.

Although individual cis-transcriptional regulatory elements associated with the chicken ovalbumin gene have been isolated and sequenced, together with short regions of flanking DNA, the entire nucleic acid sequence comprising the 5' upstream region of the ovalbumin gene has not been determined and has not been employed as a functional promoter to allow expression of a heterologous transgene.

What are still needed, however, are efficient transcription promoters that allow expression of transgenes in avian cells but with reduced positional variation.

What is also still needed is a gene expression promoter cassette that will allow expression of a transgene in the oviduct cells of an avian and efficient gene expression regardless of the chromosomal location of the expression system.

Summary of the Invention

Briefly described, the present invention relates to novel isolated and recombinant nucleic acid molecules that comprise an avian ovalbumin transcriptional regulatory region and at least one matrix attachment region element.

The isolated and recombinant nucleic acid molecules of the present invention, because of the presence of at least one matrix attachment region, are useful for reducing chromosomal positional effects on a transgene operably linked to the ovalbumin transcriptional regulatory region and transfected into a recipient cell. Isolating an approximately 195 kb region of the chicken genome that includes regions upstream of the ovalbumin locus ensures that cis-elements are also included that will

allow gene expression in a tissue-specific manner. The ovalbumin promoter region of the present invention, therefore, will allow expression of an operably linked heterologous nucleic acid insert by a transfected avian cell such as, for example, a somatic cell.

The present invention provides a novel isolated nucleic acid molecule of approximately 195 kb of the chicken genome, and truncated variants thereof, comprising a region of about 135 kb that is 5' upstream, and an approximately 45 kb region that is 3' downstream, of the ovalbumin-encoding region of the gene locus. The novel isolated chicken nucleic acid sequence includes matrix attachment regions both 5' and 3' of the ovalbumin gene and an ovalbumin transcriptional regulatory region that includes CR1 repeat elements, a proximal ovalbumin promoter. Interspersed among the elements are stretches of nucleic acid that serve at least to organize the elements in an ordered array. The novel isolated chicken genomic region also includes the ovalbumin-encoding region with a plurality of introns dispersed therein.

The present invention further provides recombinant nucleic acid molecules for operably linking an avian ovalbumin transcriptional regulatory region to a heterologous nucleic acid molecule insert encoding a polypeptide to be expressed by a transfected or transgenic cell. The heterologous nucleic acid molecule may be placed in frame with a signal peptide sequence. Translation initiation may start with the signal peptide and continue through the nucleic acid molecule to produce an expressed polypeptide having the desired amino acid sequence.

The sequence of the expressed heterologous nucleic acid insert may be optimized for codon usage by a host cell using approaches well known in the art. For example, codon usage may be optimized for an avian such as a chicken. This could be determined from the codon usage of at least one, and preferably more than one, protein expressed in a chicken cell. For example, the codon usage may be determined from the nucleic acid sequences encoding the proteins ovalbumin, lysozyme, ovomucin and ovotransferrin of chicken.

The recombinant nucleic acid molecules of the present invention may further comprise a polyadenylation signal sequence that allows transcription directed by an

ovalbumin transcriptional regulatory region to extend beyond the heterologous nucleic acid encoding a desired heterologous polypeptide and to comprise a 3' untranslated region and a polyadenylated tail. Any suitable functional polyadenylation signal sequence may be linked to the 3' end of the heterologous nucleic acid insert, including the SV40 polyadenylation signal sequence, bovine growth hormone adenylation sequence or the like.

The recombinant nucleic acid molecules of the present invention may also comprise a chicken ovalbumin 3' domain. The 3' domain can include a 3' untranslated region of the ovalbumin gene, a polyadenylation signal and at least one MAR that, in combined action with an MAR upstream of the ovalbumin transcriptional regulatory region, may reduce positional variation in gene expression in transgenic avians.

Yet another aspect of the present invention is expression vectors suitable for delivery to a recipient cell, preferably an avian cell. The expression vectors provided by the present invention may comprise an avian ovalbumin transcriptional regulatory region that can be operably linked to a nucleic acid insert encoding a polypeptide, and optionally a polyadenylation signal sequence. The expression vectors of the present invention further comprise at least one MAR element, and preferably two MARs that flank the ovalbumin transcriptional regulatory region and which can non-randomly direct the insertion of the expression vector into the genome of a recipient eukaryotic cell. The expression vector may further comprise a bacterial plasmid sequence, a viral nucleic acid sequence, or fragments or variants thereof that may allow for replication of the vector in a suitable host.

Another aspect of the present invention is methods of expressing a heterologous polypeptide in a eukaryotic cell by transfecting the cell with a recombinant nucleic molecule comprising an avian ovalbumin transcriptional regulatory region operably linked to a nucleic acid insert encoding a polypeptide desired to be expressed and, optionally, a polyadenylation signal sequence, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian ovalbumin transcriptional regulatory region.

Also within the scope of the present invention are recombinant cells, tissues and animals containing non-naturally occurring recombinant nucleic acid molecules according to the present invention as described above. In one embodiment of the present invention, the transformed cell is a chicken oviduct cell and the nucleic acid insert comprises the chicken ovalbumin transcriptional regulatory region, a nucleic acid insert encoding a human interferon α2b that is codon optimized for expression in an avian cell, and an SV40 polyadenylation sequence. In another embodiment of the present invention, the nucleic acid insert encodes the heavy and light chains of an antibody.

Additional objects and aspects of the present invention will become more apparent upon review of the detailed description set forth below when taken in conjunction with the accompanying figures, which are briefly described as follows.

Brief Description of the Figures

Fig. 1 illustrates the nucleic acid sequence SEQ ID NO: 1 of a region of the chicken genome that includes a chicken ovalbumin transcriptional regulatory region and the chicken ovalbumin gene, and matrix attachment regions 5' upstream and 3' downstream thereof.

Fig. 2 schematically illustrates the chicken genomic region having nucleic acid sequence SEQ ID NO: 1, indicating the relative positions and orientations of regions having identity with known domains.

Fig. 3 illustrates schematically the construction of an expression bacterial artificial chromosome where the insert gene of interest is under the expression control of the chicken ovalbumin promoter. Genes of interest may be inserted into the native translation start site of the ovalbumin gene. L and roman numerals, ovalbumin exons; GOI, gene of interest; start, translation start site; stop, translation stop site; pA, polyadenylation signal; E, EcoRl site.

Fig. 4 illustrates an SV40 polyadenylation signal sequence SEQ ID NO: 2.

Fig. 5 illustrates the nucleotide sequence SEQ ID NO: 3 of a human interferon α2b interferon optimized for expression in an avian cell.

Fig. 6 illustrates the reconstruction of the chicken genomic region containing the ovalbumin locus.

Detailed Description of the Preferred Embodiments

This description uses gene nomenclature accepted by the Cucurbit Genetics Cooperative as it appears in the *Cucurbit Genetics Cooperative Report* 18:85 (1995), which are incorporated herein by reference in its entirety. Using this gene nomenclature, genes are symbolized by italicized Roman letters. If a mutant gene is recessive to the normal type, then the symbol and name of the mutant gene appear in italicized lower case letters.

For convenience, definitions of certain terms employed in the specification, examples, and appended claims are collected here.

Definitions

The term "avian" as used herein refers to any species, subspecies or race of organism of the taxonomic class *ava*, such as, but not limited to chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of *Gallus gallus*, or chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Ausstralorp, Minorca, Amrox, California Gray, Italian Partidge-colored), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities. It also includes an individual avian organism in all stages of development, including embryonic and fetal stages. The term "avian" also may denote "pertaining to a bird", such as "an avian (bird) cell."

The term "nucleic acid" as used herein refers to any natural or synthetic linear and sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, such nucleic acids may be collectively referred to herein as "constructs," "plasmids," or "vectors." The term "nucleic acid" further includes modified or derivatized nucleotides and nucleosides such as, but not limited to, halogenated nucleotides such as, but not only, 5-bromouracil, and derivatised

nucleotides such as biotin-labeled nucleotides.

The term "isolated nucleic acid molecule" as used herein refers to a nucleic acid molecule with a structure not identical to a naturally occurring nucleic acid molecule and includes DNA, RNA, or derivatives or variants thereof. The term covers, but is not limited to, (a) a DNA which has the sequence of part of a naturally occurring genomic molecule but is not flanked by at least one of the coding sequences that flank that part of the molecule in the genome of the species in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic nucleic acid of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any vector or naturally occurring genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), ligase chain reaction (LCR) or chemical synthesis, or a restriction fragment; (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein, and (e) a recombinant nucleotide sequence that is part of a hybrid sequence that is not naturally occurring. Isolated nucleic acid molecules of the present invention can include, for example, natural allelic variants as well as nucleic acid molecules modified by nucleotide deletions, insertions, inversions, or substitutions such that the resulting nucleic acid molecule still essentially encodes an ovalbumin transcriptional regulatory region or a variant thereof of the present invention.

The terms "polynucleotide," "oligonucleotide," and "nucleic acid sequence" are used interchangeably herein and include, but are not limited to, coding sequences (polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory or control sequences); control sequences (e.g., translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals, transcription factor binding sites, transcription termination sequences, upstream and downstream regulatory domains, enhancers, silencers, and the like); and regulatory sequences (DNA sequences to which a transcription factor(s) binds and alters the activity of a gene's promoter either positively (induction) or negatively (repression)).

No limitation as to length or to synthetic origin are suggested by the terms described above.

As used herein the terms "peptide," "polypeptide" and "protein" refer to a polymer of amino acids in a serial array, linked through peptide bonds. A "peptide" typically is a polymer of at least two to about 30 amino acids linked in a serial array by peptide bonds. The term "polypeptide" includes proteins, protein fragments, protein analogues, oligopeptides and the like. The term "polypeptides" contemplates polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (isolated from an appropriate source such as a bird), or synthesized. The term "polypeptides" further contemplates polypeptides as defined above that include chemically modified amino acids or amino acids covalently or noncovalently linked to labeling moieties.

The term "fragment" as used herein refers to any isolated portion of the subject nucleic acid molecule constructed artificially (e.g., by chemical synthesis) or by cleaving a natural product into multiple pieces, using restriction endonucleases or mechanical shearing, or a portion of a nucleic acid synthesized by DNA polymerase, including by PCR, or any other polymerizing technique well known in the art, or expressed in a host cell by recombinant nucleic acid technology well known to one of skill in the art. The term "fragment" as used herein may also refer to an isolated portion of a polypeptide, wherein the portion of the polypeptide is cleaved from a naturally occurring polypeptide by proteolytic cleavage by at least one protease, or is a portion of the naturally occurring polypeptide synthesized by chemical or recombinant methods well known to one of skill in the art.

The terms "recombinant nucleic acid" and "recombinant DNA" as used herein refer to combinations of at least two nucleic acid sequences that are not naturally found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques. A recombinant polypeptide may be distinct from a naturally occurring

polypeptide either in its location, purity or structure. Generally, a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

The term "gene" or "genes" as used herein refers to nucleic acid sequences that encode genetic information for the synthesis of a whole RNA, a whole protein, or any portion of such whole RNA or whole protein. Genes that are not naturally part of a particular organism's genome are referred to as "foreign genes," "heterologous genes" or "exogenous genes" and genes that are naturally a part of a particular organism's genome are referred to as "endogenous genes". The term "gene product" refers to an RNA or protein that is encoded by the gene. "Endogenous gene products" are RNAs or proteins encoded by endogenous genes. "Heterologous gene products" are RNAs or proteins encoded by "foreign, heterologous or exogenous genes" and are, therefore, not naturally expressed in the cell.

The term "expressed" or "expression" as used herein refers to the transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a region of one of the two nucleic acid strands of the gene. The term "expressed" or "expression" as used herein may also refer to the translation from an RNA molecule to give a protein, a polypeptide or a portion thereof.

As used herein, the term "locus" refers to the site of a gene on a chromosome. In diploid organisms, pairs of genes control hereditary traits, each in the same position on a pair of chromosomes. These gene pairs, or alleles, may both be dominant or both be recessive in expression of that trait. In either case, the individual is said to be homozygous for the trait controlled by that gene pair. If the gene pair (alleles) consists of one dominant and one recessive trait, the individual is heterozygous for the trait controlled by the gene pair.

The term "operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Control sequences operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The control sequences need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. For example, intervening untranslated yet transcribed sequences can be present

between a promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

The term "transcription regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the transcriptional expression of the gene. Exemplary transcription regulatory sequences include enhancer elements, hormone response elements, steroid response elements, negative regulatory elements, and the like.

The term "promoter" as used herein refers to the DNA sequence that determines the site of transcription initiation by an RNA polymerase. A "promoter-proximal element" is a regulatory sequence generally within about 200 base pairs of the transcription start site.

The term "matrix attachment region" as used herein refers to a region of a eukaryotic genomic DNA that can be bound to chromosomal scaffold proteins. Matrix (scaffold) attachment regions (MARs) are generally located between transcription units such that the transcription units are within chromosomal loops. The bases of the loops are connected to the scaffold proteins through the MAR at each base. MARs and MAR-like homologs are identified as several recognizable nucleic acid sequences including, but not limited to, TG-rich spans, AT-rich regions and consensus sequences as described by Wang et al, *J. Biol. Chem.* 270:23239-23242 (1995). MARs may be identified by using suitable software such as, for example, MAR-WIZ™ (Futuresoft, Michigan, USA)

The term "internal ribosome entry sites (IRES)" as used herein refers to a region of a nucleic acid, most typically an RNA molecule, wherein eukaryotic initiation of protein synthesis occurs far downstream of the 5' end of the RNA molecule. A 43S pre-initiation complex comprising the elf2 protein bound to GTP and Met-tRNA$_i^{Met}$, the 40S ribosomal subunit, and faction elf3 and 3lf1A may bind to an "IRES" before locating an AUG start codon. An "IRES" may be used to initiate translation of a second coding region downstream of a first coding region, wherein each coding region is expressed individually, but under the initial control of a single upstream promoter. An "IRES" may be located in a eukaryotic cellular mRNA.

The term "coding region" as used herein refers to a continuous linear

arrangement of nucleotides which may be translated into a polypeptide. A full length coding region is translated into a full length protein; that is, a complete protein as would be translated in its natural state absent any post-translational modifications. A full length coding region may also include any leader protein sequence or any other region of the protein that may be excised naturally from the translated protein.

The terms "complementary", "complementarity" or "complement" as used herein refers to two nucleic acid molecules that can form specific interactions with one another to form a base-paired double helix.

The term "probe" as used herein, when referring to a nucleic acid, refers to a nucleotide sequence that can be used to anneal or hybridize with and thereby identify the presence of a complementary sequence, or a complementary sequence differing from the probe sequence but not to a degree that prevents hybridization under the hybridization stringency conditions used. The probe may be modified with labels such as, but not only, radioactive groups, biotin, and the like that are well known in the art.

The term "hybridizing under stringent conditions" as used herein refers to annealing a first nucleic acid to a second nucleic acid under stringent conditions as defined below. Stringent hybridization conditions typically permit the hybridization of nucleic acid molecules having at least 70% nucleic acid sequence complementarity with the nucleic acid molecule being used as a probe in the hybridization reaction, e.g., high temperature and/or low salt content that tend to disfavor hybridization of dissimilar nucleotide sequences. Alternatively, hybridization of the first and second nucleic acid may be conducted under reduced stringency conditions, e.g., low temperature and/or high salt content that tend to favor hybridization of dissimilar nucleotide sequences. Low stringency hybridization conditions may be followed by high stringency conditions or intermediate medium stringency conditions to increase the selectivity of the binding of the first and second nucleic acids. The hybridization conditions may further include reagents such as, but not limited to, dimethyl sulfoxide (DMSO) or formamide to disfavor still further the hybridization of dissimilar nucleotide sequences. A suitable hybridization protocol may, for example, involve hybridization in 6X SSC (wherein 1X SSC comprises 0.015 M sodium citrate and

0.15 M sodium chloride), at 65° Celsius in an aqueous solution, followed by washing with 1X SSC at 65° Celsius. Formulae to calculate appropriate hybridization and wash conditions to achieve hybridization permitting 30% or less mismatch between two nucleic acid molecules are disclosed, for example, in Meinkoth *et al.*, 1984, *Anal. Biochem.* 138: 267-284; the content of which is incorporated herein by reference in its entirety. Protocols for hybridization techniques are well known to those of skill in the art and standard molecular biology manuals may be consulted to select a suitable hybridization protocol without undue experimentation. See, for example, Sambrook *et al.*, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Press, the contents of which are herein incorporated by reference in its entirety.

Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) from about pH 7.0 to about pH 8.3 and the temperature is at least about 30° Celsius for short probes (e.g., 10 to 50 nucleotides) and at least about 60° Celcius for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37° Celsius, and a wash in 1x to 2x SSC at 50 to 55° Celsius. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37° Celsius, and a wash in 0.5x to 1x SSC at 55 to 60° Celsius. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37° Celsius, and a wash in 0.1x SSC at 60 to 65° Celsius.

The terms "percent sequence identity" as used herein refers to the degree of sequence identity between two nucleic acid sequences or two amino acid sequences as determined using the algorithm of Karlin & Attschul, 1990, *Proc. Natl. Acad. Sci.* 87: 2264-2268, modified as in Karlin & Attschul, 1993, *Proc. Natl. Acad. Sci.* 90: 5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Attschul *et al.*, 1990, *J. Mol. Biol.* Q15: 403-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12, to

obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Attschul *et al.*, 1997, *Nucl. Acids Res.* 25: 3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. Other algorithms, programs and default settings may also be suitable such as, but not only, the GCG-Sequence Analysis Package of the U.K. Human Genome Mapping Project Resource Centre that includes programs for nucleotide or amino acid sequence comparisons.

The terms "vector" or "nucleic acid vector" as used herein refer to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule (RNA or DNA) that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. The term "expression vector" as used herein refers to a nucleic acid vector that comprises a transcription regulatory region operably linked to a site wherein is, or can be, inserted, a nucleotide sequence to be transcribed and, optionally, to be expressed, for instance, but not limited to, a sequence coding at least one polypeptide.

The term "transfection" as used herein refers to the process of inserting a nucleic acid into a host cell. Many techniques are well known to those skilled in the art to facilitate transfection of a nucleic acid into an eukaryotic cell. These methods include, for instance, treating the cells with high concentrations of salt such as a calcium or magnesium salt, an electric field, detergent, or liposome mediated transfection, to render the host cell competent for the uptake of the nucleic acid molecules, and by such methods as micro-injection into a pro-nucleus, sperm-mediated and restriction-mediated integration.

The terms "recombinant cell" and "genetically transformed cell" refer to a cell comprising a combination of nucleic acid segments not found in a single cell with each other in nature. A new combination of nucleic acid segments can be introduced into an organism using a wide array of nucleic acid manipulation techniques available to those skilled in the art. A recombinant cell can be a prokaryotic cell, or a

eukaryotic cell, such as, but not limited to, an avian cell. The recombinant cell may harbor a vector that is extragenomic, i.e.that does not covalently insert into the cellular genome, including a non-nuclear (e.g. mitochondrial) genome(s). A recombinant cell may further harbor a vector or a portion thereof that is intragenomic, i.e. covalently incorporated within the genome (including non-nuclear genome(s)) of the recombinant cell.

As used herein, a "transgenic avian" is any avian, as defined above, including the chicken, in which one or more of the cells of the avian contain heterologous nucleic acid introduced by manipulation, such as by transgenic techniques. The nucleic acid may be introduced into a cell, directly or indirectly, by introduction into a precursor of the cell by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. Genetic manipulation also includes classical cross-breeding, or *in vitro* fertilization. A recombinant DNA molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The terms "chimeric animal" or "mosaic animal" are used herein to refer to animals in which the recombinant gene is found, or in which the recombinant is expressed, in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that the recombinant gene is present and/or expressed in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence that is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (*e.g.*, it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout).

The term "chromosomal positional effect" as used herein refers to the variation in the degree of gene transcription as a function of the location of the transcribed locus within the cell genome. Random transgenesis may result in a transgene being inserted at different locations in the genome so that individual cells

of a population of transgenic cells may each have at least one transgene, each at a different location and therefore each in a different genetic environment. Each cell, therefore, may express the transgene at a level specific for that particular cell and dependent upon the immediate genetic environment of the transgene. In a transgenic animal, as a consequence, different tissues may exhibit different levels of transgene expression. The term "reduced chromosomal positioning effect" as used herein refers to a decreased intercellular variation in the level of gene transcription because of a reduction in the number of sites of insertion of a heterologous nucleic acid molecule into the genome of a recipient cell. Consequently, a reduced chromosomal positioning effect provides a more uniform population of genetically transformed cells with respect to transgene insertion sites in the cellular genomes. In transgenic animals, different tissues may exhibit reduced variability in the levels of transgene expression.

The term "cytokine" as used herein refers to any secreted polypeptide that affects a function of cells and modulates an interaction between cells in the immune, inflammatory or hematopoietic response. A cytokine includes, but is not limited to, monokines and lymphokines. Examples of cytokines include, but are not limited to, interferon α2b, Interleukin-1 (IL-1), Interleukin-6 (IL-6), Interleukin-8 (IL-8), Tumor Necrosis Factor-α (TNF-α.) and Tumor Necrosis Factor β (TNF-β.).

The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof. Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The term "immunoglobulin polypeptide" as used herein refers to a constituent polypeptide of an antibody or a polypeptide derived therefrom. An "immunological polypeptide" may be, but is not limited to, an immunological heavy or light chain and may include a variable region, a diversity region, joining region and a constant region or any combination, variant or truncated form thereof. The term "immunological polypeptides" further includes single-chain antibodies comprised of, but not limited

to, an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region and optionally a peptide linker.

Techniques useful for isolating and characterizing the nucleic acids and proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals may be consulted to select suitable protocols without undue experimentation. See, for example, Sambrook *et al*, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety.

Abbreviations

Abbreviations used in the present specification include the following: aa, amino acid(s); bp, base pair(s); kb, kilobase; cDNA, DNA complementary to RNA; SSC, sodium chloride-sodium citrate; DMSO, dimethyl sulfoxide; MAR, matrix attachment region; CPE, chromosomal positioning effect; BAC, bacterial artificial chromosome; YAC, yeast artificial chromosome.

The present invention provides novel isolated and recombinant nucleic acid molecules comprising an avian ovalbumin transcriptional regulatory region and at least one MAR element, which are useful as vectors for inserting a heterologous nucleic acid molecule into the genome of a recipient avian cell. The novel isolated nucleic acid molecules of the present invention are particularly useful for directing the incorporation of a heterologous nucleic acid that is under transcriptional regulation of an avian ovalbumin gene promoter, into the genome of a recipient avian cell while reducing or avoiding chromosomal positioning effects that would otherwise result from randomly distributed insertions of the heterologous nucleic acid molecule into the recipient avian genome. The present invention further provides methods of delivering a heterologous nucleic acid under the transcriptional regulation of an avian ovalbumin transcriptional regulatory region, to an avian cell, whereby the heterologous nucleic acid desired to be expressed under the associated avian ovalbumin gene transcriptional regulatory element can be integrated into an avian cell genome. As well as providing recombinant nucleic acids, vectors and derivatives

thereof, the present invention provides transfected and transgenic avian cells and birds derived therefrom that are capable of producing a heterologous polypeptide in the serum or the white of a laid egg.

Nucleic acids comprising the chicken ovalbumin gene and 5' and 3' MAR elements

The novel isolated and recombinant nucleic acid molecules of the present invention comprise the chicken ovalbumin gene comprising transcriptional regulatory elements positioned 5′ upstream of the ovalbumin-encoding region of the native chicken ovalbumin locus and which are necessary for the regulated expression of a downstream polypeptide-encoding nucleic acid, and at least one MAR element.

The inclusion of a MAR element, and preferably at least two MARs, in the same nucleic acid and flanking the ovalbumin gene region, may confer positional independence to a transfected gene operably linked to the ovalbumin transcriptional regulatory region. While not wishing to be bound by any one theory, it is believed that the 5' and 3' MARs of a transfected nucleic acid molecule of the present invention restrict the number of possible transgene insertion sites within the genome of the recipient avian cell, thereby reducing chromosomal positioning effects upon transcription levels. Thus the isolated novel nucleic acid molecules of the present invention are useful for reducing the chromosomal positional effects exerted on heterologous transgene expression. The heterologous transgene will be operably linked to the ovalbumin transcriptional regulatory region within a novel recombinant nucleic acid molecule transfected into a recipient avian cell. Included in the nucleic acid molecules of the present invention are a region of the avian genome encompassing a MAR upstream of the ovalbumin locus and cis-regulatory elements that may allow gene expression in a tissue-specific manner. The ovalbumin promoter region of the novel nucleic acid molecules is especially useful for directing expression of an operably linked heterologous nucleic acid in a transfected avian cell such as an avian oviduct cell.

Also within the scope of the present invention that nucleic acid molecules further comprising a region of the chicken ovalbumin locus that is 3' of the ovalbumin-encoding region, or of a nucleic acid insert encoding a heterologous

present invention includes at least one nucleic acid sequence encoding a 3' MAR element which may cooperate with a 5' MAR to limit the number of sites of insertion into the genome of an avian cell of a transfected nucleic acid molecule. In either event, the directed insertion induced by one or more MARs can reduce or eliminate chromosomal positioning effects, resulting in a more uniform level of gene expression of the heterologous nucleic acid insert in a population of genetically transformed cells.

(a) *Isolated nucleic acid encompassing the chicken ovalbumin gene*

One aspect of the present invention, therefore, is a nucleic acid molecule isolated from the genome of a chicken and comprising a proximal ovalbumin promoter suitable for directing transcription of regulation of a transcript encoding ovalbumin, and 5' and 3' MAR elements flanking the ovalbumin gene region.

BACs 120 and 77 (ATCC Accession Nos. _____) containing overlapping regions of the chicken genome, were sequenced and compiled as the contiguous sequence SEQ ID NO: 1. BAC 120 includes the sequence from nucleotide position 1 to position 157354 of SEQ ID NO: 1. The sequence of BAC 77 begins at nucleotide position 157355 of SEQ ID NO: 1 to position 195102. The nucleic sequence of the 195,102 bp chicken genomic region SEQ ID NO: 1 (GenBank Accession No. _____) is shown in Fig. 1. A schematic showing identifiable domains within SEQ ID NO: 1 that have sequence identity or homology to known domain families or previously identified genes mainly identified using BLAST, GenScan and MARWIZ software is shown in Fig. 2. BAC 26, constructed as described in Example 1 below and containing the entire nucleic acid insert SEQ ID NO: 1 less about 11.5 kb at the extreme 5' end, was deposited with American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110, as ATCC No. PTA-5548 on September 24, 2003, under the conditions set forth in the Budapest Treaty.

The nucleic acid molecule SEQ ID NO: 1 of the present invention has at least four MAR elements. One MAR element is 5' upstream of the ovalbumin gene, between about nucleotide positions 41701 and 41900. MAR-like elements are also between nucleotide positions 56001-56201, 56501-56901, 58401-58701, 76251-

between nucleotide positions 56001-56201, 56501-56901, 58401-58701, 76251-76451 and 80151-80451. Another MAR element is between about 96401-96800. MAR elements located 3' downstream of the ovalbumin gene are at nucleotide positions about 144651-144850, about 150601-151600, about 156681-157181, about 157081-15781, about 163701-164100, about 186201-186590 and about 190101-190800 of SEQ ID NO: 1. The chicken ovalbumin gene ATG start codon is at nucleotide position 133372.

Also dispersed along the nucleic acid molecule represented by SEQ ID NO: 1 are other identifiable domains listed, for example, in Table 2 below, including several serpin- or serpin-like encoding genes, cis transcription regulatory elements of both the serpin-like and ovalbumin genes, and at least two other, putatively functional genes, X and Z. Between the various domains, genes or other elements are stretches of nucleotides, the functions of which may serve to maintain the position and configuration of the elements relative to each other.

The isolated nucleic acids of the present invention and derivatives and truncated variants thereof may be incorporated into a vector, such as a bacterial or yeast artificial chromosome. The BAC cloning system (Shizuya *et al*, *Proc. Natl. Acad. Sci (U.S.A.)*, 89:8794:8797, (1992) has been developed to stably maintain large fragments of genomic DNA (100-300 kb) in *E. coli*. An exemplary BAC vector consists of the pBeloBAC11 vector that has been described by Kim *et al*, *Genomics*, 34:213:218 (1996). Genomic DNA can be partially digested, for example, using enzymes that permit ligation into either the BamH I or Hind III sites in the vector. It is contemplated that any suitable restriction sites may be used that are useful for incorporating genomic DNA into a selected BAC vector. Flanking these cloning sites are T7 and SP6 RNA polymerase transcription initiation sites that can be used to generate end probes by either RNA transcription or PCR methods. BAC DNA is purified from the host cell as a supercoiled circle. Converting these circular molecules into a linear form precedes both size determination and introduction of the BACs into recipient avian cells. A suitable cloning site may be flanked, for example, by two Not I restriction sites, permitting cloned segments to be excised from the vector by Not I digestion. Alternatively, the BAC vector may be linearized by

treatment with the commercially available enzyme lambda terminase that leads to the cleavage at a cosN site. However, this cleavage method results in a full length BAC clone containing both the insert DNA and the BAC vector sequences.

One embodiment of the novel isolated nucleic acid molecules of the present invention, therefore, is an isolated chicken nucleic acid molecule encoding an ovalbumin transcriptional regulatory region and a 5' MAR. In one embodiment of the present invention, the novel isolated nucleic acid molecule further comprises a 3' MAR downstream of the ovalbumin gene. The isolated nucleic acid molecules of the present invention may also include nucleic acid elements such as, but not limited to, a transcription enhancer element, a negative regulator element, a hormone responsive element, an avian CR1 repeat element that together may constitute, in whole or in part, the ovalbumin transcriptional regulatory region, a proximal ovalbumin promoter and a signal peptide-encoding region. There are also stretches of nucleic acid between these constituent elements that organize the various elements into an ordered linear array. While the constituent elements of the ovalbumin transcriptional control region are preferably ordered as in sequence SEQ ID NO: 1, it is within the scope of the present invention for the cis-elements of the ovalbumin transcriptional regulatory region to be in any linear arrangement that will allow the formation of a transcript comprising the nucleotide sequence, or its complement, of a nucleic acid insert operably linked to the ovalbumin transcriptional regulatory region.

The novel isolated nucleic acid molecules of the present invention allow one skilled in the art to, for example, (a) make copies of those nucleic acid molecules by procedures such as, but not limited to, insertion into a cell for replication by the cell, by chemical synthesis or by procedures such as PCR or LCR, (b) obtain nucleic acid molecules which include at least a portion of such nucleic acid molecules, including full-length genes, full-length coding regions, transcriptional regulatory sequences, truncated coding regions and the like, (c) identify and obtain ovalbumin transcriptional regulatory region homologs found in other avian species such as, but not limited to, turkey, duck, goose, quail, pheasant, parrot, finch, ratites including ostrich, emu and cassowary and, (d) to obtain isolated nucleic acids capable of hybridizing to an avian ovalbumin transcriptional regulatory region nucleic acid and

of being used as a probe to detect the presence of nucleic acid-related sequences by complementation between the probe and the target nucleic acid.

Such nucleic acid homologs can be obtained in a variety of ways including using traditional cloning techniques to screen appropriate libraries, amplifying appropriate libraries or DNA using oligonucleotide primers derived from the novel nucleic acid molecules of the present invention in a polymerase chain reaction or other amplification method, and screening public and/or private databases containing genetic sequences using nucleic acid sequences of the present invention to identify targets. Examples of preferred libraries to screen, or from which to amplify nucleic acid molecules, include but are not limited to avian BAC libraries, genomic DNA libraries, and cDNA libraries. Similarly, preferred sequence databases useful for screening to identify sequences in other species homologous to chicken ovalbumin transcriptional regulatory region include, but are not limited to, GenBank and the mammalian Gene Index database of The Institute of Genomics Research (TIGR).

Nucleotides used to construct the nucleic acids of the present invention can be labeled to provide a signal as a means of detection, using conventional labeling technologies such. as radioactive labels, fluorescent compounds, enzymes and chemiluminescent moieties. Methods useful in selecting appropriate labels and binding protocols for binding the labels to the synthetic nucleotides are well known to those of skill in the art.

In one embodiment of the isolated nucleic acid molecule according to the present invention, the nucleic acid is isolated from a chicken.

In other embodiments of the isolated nucleic acid molecule according to the present invention, the nucleic acid molecule comprises a nucleotide sequence having at least 80% identity, at least 95% identity or at least 99% identity to the nucleotide sequence according to SEQ ID NO: 1, or the complement thereof.

In other embodiments, the isolated nucleic acid molecule of the invention comprises the nucleotide sequence according to SEQ ID NO: 1 or has the nucleotide sequence according to SEQ ID NO: 1. In another embodiment, the isolated nucleic acid molecule can be an allelic variant of SEQ ID NO: 1.

(b) *Fragments and Variants of SEQ ID NO: 1*

Fragments of the isolated nucleic acid molecules of the present invention also are within the scope of the present invention. As used herein, a fragment of a nucleic acid molecule refers to a nucleotide sequence having fewer nucleotides than the nucleotide sequence SEQ ID NO: 1 but which includes a nucleic acid sequence of the ovalbumin transcriptional regulatory region able to direct and regulate transcription of a nucleic acid, and at least one MAR element.

The isolated nucleic acid molecule having the sequence SEQ ID NO: 1 may be reduced in size by truncating regions that do not affect the expression of a heterologous nucleic acid placed under the transcriptional control of the ovalbumin transcription regulatory region. A truncated variant of the nucleic acid molecule of the present invention is understood to be any variant of SEQ ID NO: 1 less nucleotides at either the 5' and/or the 3' end of SEQ ID NO: 1. For example, it is contemplated that any of the nucleotides from positions 1- about 40500 may be individually, in part, or in total, deleted from the variant nucleic acid molecule. Similarly, nucleotides from positions about 151700, 164200, 186690 or 190900 to 195101 of the nucleic acid molecule having sequence SEQ ID NO: 1 may be removed, retaining 1, 2, 3, 4 or more of the MAR or MAR-like elements respectively located 3' at the chicken ovalbumin gene. Useful truncated variants of SEQ ID NO: 1, therefore, include, but are not limited to, from base position about 41000 to about 191500, to about 187000, to about 164500, to about 152000, or to about 145500 and from base position about 96000 to about 191500, to about 187000, to about 164500, to about 152000 or to about 145500. Other useful truncated variants of SEQ ID NO: 1 include regions from nucleotide positions about 56000, about 58350, about 76200 and about 80000 to about 191500, to about 187000, to about 164500, to about 152000 or to about 145500.

Therefore, the invention encompasses nucleic acid molecules which do not include regions that do not contribute to the desired functionality of inserting a heterologous nucleic acid into an avian genome with reduced or no chromosomal positioning effect. The region 5' upstream of the MAR located at nucleotide positions 41701-41900 of SEQ ID NO: 1, may be deleted to give a truncated variant of SEQ ID

NO: 1. For example, the approximately 11.5 kb region extending from nucleotide position 1 of SEQ ID NO: 1 not present in BAC 26 may be deleted. Likewise, it is contemplated that other regions of SEQ ID NO: 1 as listed in Table 2, such as encoding the serpin-like proteins, may be selectively deleted.

Recombinant nucleic acids

Another aspect of the present invention is recombinant nucleic acid molecules comprising at least one, and preferably at least two, avian MARs and an avian ovalbumin transcription regulatory region, including the proximal promoter thereof. The recombinant nucleic acid molecules of the present invention are particularly useful for delivering a desired heterologous nucleic acid to a recipient avian cell while reducing chromosomal positional effects upon transcription from the integrated heterologous nucleic acid. It is contemplated that regions of SEQ ID NO: 1 may be omitted from the recombinant nucleic acid molecules of the present invention without substantially affecting the reduction in the CPE compared to a similar nucleic acid molecule not including MAR elements. For example, one or more of the serpin-encoding regions of SEQ ID NO: 1 listed in Table 2, below, may not be included.

The present invention, therefore, provides recombinant nucleic acid molecules that comprise at least one avian MAR and an avian ovalbumin transcription regulatory region optionally operably linked in a linear array to a selected heterologous or endogenous polypeptide-encoding nucleic acid insert, and which may express the nucleic acid insert when transfected to a suitable host cell, preferably an avian cell.

The nucleic acid insert, such as a heterologous nucleic acid can be operably linked 3' downstream of the ovalbumin proximal promoter and is thereby expressed as an RNA transcript by a transfected recipient cell. The heterologous nucleic acid may be inserted into the recombinant nucleic acid of the present invention 3' downstream of a region encoding a peptide leader region so that a heterologous polypeptide encoded by the inserted nucleic acid may include this leader region. It is within the scope of the present invention for the recombinant nucleic acid to have the nucleic acid insert encoding the desired polypeptide to be operably inserted into the ovalbumin coding region, or operably replacing the ovalbumin coding region in whole

or in part. The generation of BACs comprising a heterologous nucleic acid under the transcriptional control of the ovalbumin gene control region according to the present invention are described in Examples 2 and 3, below.

To increase the efficiency of expression of the heterologous nucleic acid insert, a polyadenylation signal region may be included at the 3' end of the inserted nucleic acid to allow the transcript directed by the novel ovalbumin transcriptional regulatory region to proceed beyond the nucleic acid insert encoding a selected polypeptide thereby providing a transcript further comprising a 3' untranslated region and a polyadenylated tail. Any suitable functional polyadenylation signal sequence may be linked to the 3' end of the nucleic acid insert including, for example, the SV40 polyadenylation signal sequence SEQ ID NO: 2 as shown in Fig. 4, bovine growth hormone adenylation sequence or the like. It is further anticipated that the recombinant nucleic acid molecules of the present invention may comprise the chicken ovalbumin 3' domain, or a variant thereof. The ovalbumin 3' domain may comprise the ovalbumin 3' untranslated region, an ovalbumin gene polyadenylation sequence and at least one of the 3' MAR elements identified downstream of the ovalbumin-encoding region of SEQ ID NO: 1. If the heterologous nucleic acid is inserted within the ovalbumin encoding region, and in-phase with the ovalbumin gene, the polyadenylation signal region of the ovalbumin gene may be used.

In one embodiment of the recombinant nucleic acid molecule according to the present invention, the recombinant nucleic acid molecule comprises the nucleotide sequence according to SEQ ID NO: 1, or the complement thereof.

Another aspect of the present invention is a recombinant DNA molecule comprising a MAR element and an avian ovalbumin transcriptional regulatory region. In one embodiment, the ovalbumin transcriptional regulatory region is operably linked in linear array to a nucleic acid insert encoding a polypeptide sought to be expressed, and a polyadenylation signal sequence optionally operably linked thereto. It is contemplated that when the recombinant nucleic acid molecule is to be delivered to a recipient avian cell for expression therein, the sequence of the inserted heterologous nucleic acid sequence may be modified so that the codons thereof are optimized for the codon usage of the recipient species as described below. In a

preferred embodiment, a MAR element is located 5' upstream of the ovalbumin transcriptional regulatory region. Suitable MAR elements, for example, are located at about nucleotide positions about 41701-41800 and about 96401-96800 of sequence SEQ ID NO: 1.

In one embodiment of the present invention, the recombinant nucleic acid molecule comprises the nucleotide sequence from nucleotides position about 40750 to 195101 of SEQ ID NO: 1. Various embodiments of the recombinant nucleic acid molecules of the present invention comprise a 5' MAR and/or a 3' MAR, and the ovalbumin transcriptional regulatory region. In one embodiment, the recombinant nucleic acid further comprises a T gene. In others, the recombinant nucleic acid further comprises at least one nucleic acid region selected from the group consisting of the U serpin gene, a V serpin gene, an X gene, a Y gene and a Z serpin derived from SEQ ID NO: 1.

Another embodiment of the recombinant nucleic acid molecules further comprises at least one avian MAR 3' downstream of the nucleic acid insert. Suitable MAR elements for inclusion 3' downstream of the ovalbumin transcriptional regulatory region of the recombinant construct of the present invention are found at nucleotide positions about 144651-144850, about 150800-151600, about 163701-164100, about 186201-186590 and about 190101-190800 of sequence SEQ ID NO: 1. In one embodiment of the recombinant nucleic acid molecules, the ovalbumin transcriptional regulatory region, the avian 5' MAR, and the avian 3' MAR are independently capable of hybridizing under high stringency conditions to the nucleic acid sequence according to SEQ ID NO: 1, or the complement thereof.

In various embodiments of the present invention, the recombinant nucleic acid molecule is inserted into a vector such as, but not limited to, a plasmid or viral vector.

Other embodiments of the recombinant nucleic acid molecules further comprise a plasmid or viral origin of replication. In one embodiment, the recombinant nucleic acid molecule is a bacterial or yeast artificial chromosome.

Yet another embodiment of the recombinant nucleic acid molecule according to the present invention, therefore, is a recombinant nucleic acid molecule comprising an avian ovalbumin transcription regulatory region, an avian 5' MAR, a heterologous

nucleic acid encoding a heterologous polypeptide desired to be expressed by a recipient genetically modified cell, a polyadenylation signal sequence, and an avian 3' MAR, wherein the avian ovalbumin transcription regulatory region, 5' MAR, and the 3' MAR each independently hybridizes under high stringency conditions to the nucleic acid sequence SEQ ID NO: 1, or a complement thereof.

Polypeptide expression under the control of an avian ovalbumin promoter

Another aspect of the present invention of the novel isolated ovalbumin transcriptional regulatory region is increasing the amount of a heterologous protein present in a bird (especially the chicken) by gene transfer. Typically, a heterologous polypeptide-encoding nucleic acid insert transferred into the recipient animal host will be operably linked with the ovalbumin transcriptional regulatory region to allow the cell to initiate and continue production of the genetic product protein. A recombinant DNA molecule of the present invention can be transferred into the extra-chromosomal or genomic DNA of the host.

A useful application of the novel isolated and recombinant nucleic acid molecules of the present invention is to increase the amount of a heterologous protein present in a bird, (especially the chicken) by gene transfer. Typically, a heterologous polypeptide-encoding nucleic acid insert transferred into the recipient bird host or an isolated cell or cell-line from the bird will be operably linked with the ovalbumin transcriptional regulatory region to allow the cell to initiate and continue production of the genetic protein product.

The isolated nucleic acid molecule SEQ ID NO: 1 is useful for inserting therein a heterologous nucleic acid that is desired to be expressed as a transcript or, ultimately, as a polypeptide. A heterologous nucleic acid may be operably linked to the proximal promoter region of the ovalbumin gene at any position 3' downstream of the promoter that allows transcription from the heterologous nucleic acid and synthesis of the desired encoded peptide. Some, or all, of the ovalbumin-encoding region of the isolated or recombinant nucleic acids of the present invention may be replaced by a heterologous nucleic acid to be expressed under the transcriptional control of upstream ovalbumin gene control region. The heterologous nucleic acid may be inserted into the isolated or recombinant nucleic acids of the present invention

so that the expressed amino acid sequences derived from the ovalbumin may be linked to the expressed heterologous protein either at the N-terminus or C-terminus thereof.

Any of the vectors of the present invention may also optionally include a sequence encoding a signal peptide that directs secretion of the protein expressed by the vector from the transgenic cells, for instance, from tubular gland cells of the oviduct. This aspect of the invention effectively broadens the spectrum of exogenous proteins that may be deposited in avian eggs using the methods of the invention. Where an exogenous protein would not otherwise be secreted, the vector bearing the coding sequence is modified to comprise, for instance, about 60 bp encoding a signal peptide. The DNA sequence encoding the signal peptide is inserted in the vector such that the signal peptide is located at the N-terminus of the protein encoded by the vector.

The expression vectors of the present invention comprise avian ovalbumin transcriptional regulatory regions that can direct expression of either fusion or non-fusion proteins. With fusion vectors, a number of amino acids are usually added to the desired expressed target gene sequence such as, but not limited to, a protein sequence for thioredoxin. A proteolytic cleavage site may further be introduced at a site between the target recombinant protein and the fusion sequence. Additionally, a region of amino acids such as a polymeric histidine region may be introduced to allow binding of the fusion protein to metallic ions such as nickel bonded to a solid support, for purification of the fusion protein. Once the fusion protein has been purified, the cleavage site allows the target recombinant protein to be separated from the fusion sequence. Enzymes suitable for use in cleaving the proteolytic cleavage site include, but are not limited to, Factor Xa and thrombin. Fusion expression vectors that may be useful in the present invention include pGex (Amrad Corp., Melbourne, Australia), pRIT5 (Pharmacia, Piscataway, NJ) and pMAL (New England Biolabs, Beverly, MA), that fuse glutathione S-transferase, protein A, or maltose E binding protein, respectively, to the target recombinant protein.

The present invention further relates to nucleic acid vectors and transgenes derived therefrom that incorporate polypeptide-encoding regions, wherein a first

polypeptide-encoding region is operatively linked to an avian ovalbumin promoter and a second polypeptide-encoding region is operatively linked to an Internal Ribosome Entry Sequence (IRES). It is contemplated that the first polypeptide-encoding region, the IRES and the second polypeptide-encoding region of a recombinant DNA of the present invention may be arranged linearly, with the IRES operably positioned immediately 5' of the second polypeptide-encoding region. This nucleic acid construct, when inserted into the genome of a bird and expressed therein, will generate individual polypeptides that may be post-translationally modified and combined in the white of a hard-shell bird egg. Alternatively, the expressed polypeptides may be isolated from an avian egg and combined *in vitro*.

Expression of a heterologous nucleic acid by a recombinant expression vector according to the present invention can be obtained using eukaryotic host cells, preferably avian cells, more preferably chicken cells, and still more preferably chicken oviduct cells, especially tubular gland cells. The use of eukaryotic host cells permit partial or complete post-translational modification such as, but not only, glycosylation and/or the formation of the relevant inter- or intra-chain disulfide bonds. Examples of vectors useful for expression in the chicken *Gallus gallus* include pYepSecl as in Baldari *et al.*, E.M.B.O.J., 6, 229-234 (1987) and pYES2 (Invitrogen Corp., San Diego, CA), incorporated herein by reference in their entireties.

One aspect of the present invention is methods of delivering a novel nucleic acid molecule of the present invention to the cytoplasm of an avian cell having a nucleus, thereby generating a transfected and genetically transformed avian cell. Such incorporation can be carried out by the various forms of transfection, depending upon the vector/host cell system. It is contemplated that the incorporation of recombinant nucleic acid molecules of the present invention into a recipient cell may be by any suitable method such as, but not limited to, viral transfer, electroporation, gene gun insertion, sperm-mediated transfer to an ovum, microinjection and the like.

In the various embodiments of these methods, the avian cell may be a chicken cell or a quail cell. In some embodiments of the methods of the present invention, the avian cell is within oviductal tissue of a bird, an isolated oviduct cell or primary cell

line, or a sustainable oviduct cell line. Preferably, the oviduct cells are tubular gland cells.

Heterologous polypeptide can be produced by transfected cells of the invention *in vitro*, i.e., in tissue culture outside the body of a living animal. Alternatively, the nucleic acids of the present invention may be delivered to an animal such as a chicken, whereupon the nucleic acid may enter cells and be expressed therein. It is anticipated that the nucleic acids of the present invention may integrate into the genome of the recipient cells and then express the encoded, typically heterologous, polypeptide therein. Preferably, a heterologous nucleic acid is delivered to oviduct cells within a chicken for synthesis of the desired polypeptide and its deposition in the white of an egg.

Another aspect of the present invention is a eukaryotic cell transfected with an expression vector according to the present invention and described above. For example, in one embodiment, the transformed cell can be a chicken oviduct cell or cell line, including a sustainable cell line, and the transfected nucleic acid insert comprises the chicken ovalbumin transcriptional regulatory region, a 5' MAR and/or a 3' MAR, a nucleic acid insert encoding a human interferon α2b and codon optimized for expression in an avian cell, and an SV40 polyadenylation sequence. In another example, the nucleic acid insert encodes an immunoglobulin heavy chain and a second chain under the transcriptional control of an IRES.

The transfected cell according to the present invention may be transiently transfected, whereby the transfected recombinant nucleic cid, such as DNA, or expression vector may not be integrated into the genomic nucleic acid. However, the transfected recombinant DNA or expression vector may be stably integrated into the genomic DNA of the recipient cell, thereby replicating with the cell so that each daughter cell receives a copy of the transfected nucleic acid. When the recombinant DNA or expression vector of the present invention is integrated into the genomic DNA of the recipient cell so that the cell is genetically transformed, it is anticipated that the MAR element(s) of the integrated nucleic acid will direct integration a limited number of integration site within the target genome, thereby producing a

population of cells more uniform with regard to the level of expression of the heterologous nucleic acid.

The the present invention also includes a transgenic bird producing a heterologous protein expressed from a transfected nucleic acid according to the present invention. The transgenic bird is selected from a turkey, duck, goose, quail, pheasant, ratite, an ornamental bird or a feral bird. In a preferred embodiment, the avian is a chicken and the heterologous protein produced under the transcriptional control of the avian ovalbumin transcriptional regulatory region according to the present invention is produced in the white of an egg.

Viral host cell transformation

Nucleic acid sequences or derivative or truncated variants thereof, may be introduced into viruses such as an adenovirus or vaccinia virus. Methods for making a viral recombinant vector useful for expressing a protein under the control of the ovalbumin promoter are analogous to the methods disclosed in U.S. Patent Nos. 4,603,112; 4,769,330; 5,174,993; 5,505,941; 5,338,683; 5,494,807; 4,722,848; Paoletti, E., 1996, *Proc. Natl. Acad. Sci.* 93: 11349-11353; Moss, 1996, *Proc. Natl. Acad. Sci.* 93: 11341-11348; Roizman, 1996, *Proc. Natl. Acad. Sci.* 93: 11307-11302; Frolov *et al.*, 1996, *Proc. Natl. Acad. Sci.* 93: 11371-11377; Grunhaus *et al.*, 1993, *Seminars in Virology* 3: 237-252 and U.S. Patent Nos. 5,591,639; 5,589,466; and 5,580,859 relating to DNA expression vectors, *inter alia,* the contents of which are incorporated herein by reference in their entireties.

Retrovirus vectors and adeno-associated virus vectors provide efficient systems of delivery of genes into cells, and the transferred nucleic acids may be stably integrated into the chromosomal DNA of the host. Protocols for producing recombinant retroviruses and for infecting cells *in vitro* or *in vivo* with such viruses can be found in Ausubel *et al.*, 1989, *Current Protocols in Molecular Biology* §§ 9.10-9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are well known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include psiCrip, psiCre, psi2 and psiAm.

Furthermore, it is possible to limit the infection spectrum of retroviruses and consequently of retroviral-based vectors, by modifying the viral packaging proteins on the surface of the viral particle (see, for example PCT publications WO 93/25234, WO 94/06920, and WO 94/11524). Roux *et al.*, 1898, *Proc. Natl. Acad. Sci.* 86:9079-9083; Julan *et al.*, 1992, *J. Gen. Virol.* 73:3251-3255; and Goud *et al.*, 1983, *Virology* 163:251-254); Neda *et al.*, 1991, *J. Biol. Chem.* 266:14143-14146), which are incorporated herein by reference in their entireties.

One retrovirus for randomly introducing a transgene into the avian genome is a replication-deficient ALV retrovirus. To produce an appropriate ALV retroviral vector, a pNLB vector may be modified by inserting a region comprising at least part of the ovalbumin transcriptional regulatory region, a MAR element and one or more exogenous genes between the 5' and 3' long terminal repeats (LTRs) of the retrovirus genome. Any coding sequence placed in-frame and downstream of the ovalbumin promoter will be expressed at high levels and especially in the tubular gland cells of the oviduct magnum because the ovalbumin promoter drives the high level of expression of the ovalbumin protein and is only active in the oviduct tubular gland cells.

Another viral gene delivery system useful in the present invention utilizes adenovirus-derived vectors (see, for example, Berkner *et al.*, 1988, *BioTechniques* 6:616-629; Rosenfeld *et al.*, 1991, *Science* 252:431-434; and Rosenfeld *et al.*, 1992, *Cell* 68:143-155), incorporated herein by reference in their entireties. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are well known to those skilled in the art. Introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situations where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA).

Yet another viral vector system is the adeno-associated virus (AAV). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. In the present invention, at least part of the heterologous nucleic acid will include an operable region of the avian ovalbumin transcriptional regulatory region and a MAR

element. An AAV vector such as that described in Tratschin *et al.*, 1985, *Mol. Cell. Biol.* 5:3251-3260, can be used to introduce DNA into cells.

Other viral vector systems that may have application in the methods according to the present invention have been derived from, but are not limited to, herpes viruses, vaccinia viruses, avian leucosis viruses and several RNA viruses.

Non-viral expression vectors

Most non-viral methods of gene transfer rely on normal mechanisms used by eukaryotic cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject ovalbumin transcriptional regulatory region and operably linked polypeptide-encoding nucleic acid by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

In a representative embodiment, a nucleic acid comprising the novel recombinant nucleic acids of the present invention can be entrapped in liposomes bearing positive charges on their surface (*e.g.*, lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno *et al.*, 1992, *NO Shinkei Geka* 20: 547-551; PCT publication WO91/06309; Japanese patent application 1047381; and European patent publication EP-A-43075, all of which are incorporated herein by reference in their entireties).

In similar fashion, the gene delivery system comprises an antibody or cell surface ligand that is cross-linked with a gene binding agent such as polylysine (see, for example, PCT publications WO93/04701, WO92/22635, WO92/20316, WO92/19749, and WO92/06180, all of which are incorporated herein by reference in their entireties). It will also be appreciated that effective delivery of the subject nucleic acid constructs via receptor-mediated endocytosis can be improved using agents which enhance escape of genes from the endosomal structures. For instance, whole adenovirus or fusogenic peptides of the influenza HA gene product can be used as part of the delivery system to induce efficient disruption of DNA-containing endosomes (Mulligan *et al.*, 1993, *Science* 260-926; Wagner *et al.*, 1992, *Proc. Natl. Acad. Sci.* 89:7934-7938; and Christiano *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2122-

2126, all of which are incorporated herein by reference in their entireties). It is further contemplated that a recombinant DNA molecule of the present invention may be delivered to a recipient host cell by other non-viral methods including by gene gun, microinjection, sperm-mediated transfer, or the like.

Another aspect of the present invention is a method of expressing a heterologous polypeptide in a eukaryotic cell by transfecting a cell with a recombinant nucleic acid molecule of the invention, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian ovalbumin transcriptional regulatory region.

In one embodiment of this aspect, the nucleic acid molecule is integrated into the genome of the recipient avian cell. In some embodiments the recipient avian oviduct cell is a chicken cell, preferably a chicken oviduct cell, more preferably an oviduct tubular gland cell.

The protein of the present invention may be produced in purified form by any known conventional techniques. For example, chicken cells, an egg or an egg white may be homogenized and centrifuged. The supernatant may then be subjected to sequential ammonium sulfate precipitation and heat treatment. The fraction containing the protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If necessary, the protein fraction may be further purified by HPLC or other methods well known in the art of protein purification.

Expression of heterologous multimeric proteins by transfected avian cells

The present invention provides methods for the production of a multimeric protein by an avian cell, comprising the step of culturing an avian cell transfected with a first expression vector and, optionally, a second expression vector; the expression vectors may each have a transcription unit comprising a nucleotide sequence encoding a first heterologous polypeptide, a transcription promoter, and a transcriptional terminator operatively linked to the nucleotide sequence encoding a second heterologous polypeptide, such that the cultured avian cell produces a multimeric protein comprising the first and second heterologous polypeptides.

The isolated nucleic acids and recombinant nucleic acid constructs derived therefrom of the present invention are useful to express nucleic acid sequences of polypeptides that are optimized for expression in avian cells, and derivatives and fragments thereof. Such derivatives include, for instance, polypeptides with conservative amino acid replacements, that is, those within a family of amino acids that are related in their side chains (commonly known as acidic, basic, nonpolar, and uncharged polar amino acids). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids and other groupings are known in the art (see, for example, "Biochemistry", 2nd ed, L. Stryer, ed., WH Freeman and Co.,1981). Peptides in which more than one replacement has taken place can readily be tested for activity in the same manner as derivatives with a single replacement, using conventional polypeptide activity assays (e.g. for enzymatic or ligand binding activities).

Regarding codon optimization, for example, if the recombinant DNA is transfected into a recipient chicken cell, the sequence of the nucleic acid insert to be expressed is optimized for chicken codon usage. This may be determined from the codon usage of at least one, and preferably more than one, protein expressed in a chicken cell according to well known principles. For example, in the chicken the codon usage may be determined from the nucleic acid sequences encoding the proteins lysozyme, ovalbumin, ovomucin and ovotransferrin of chicken. Optimization of the sequence for codon usage elevates the level of translation in avian eggs.

One embodiment of the recombinant nucleic acid of the present invention, comprises an insert encodes the human interferon α2b polypeptide. The exemplary nucleic acid sequence SEQ ID NO: 3 (Fig. 5) encodes the polypeptide human interferon α2b in accordance with avian cell codon usage, as determined from the nucleotide sequences encoding chicken ovomucin, ovalbumin, ovotransferrin and lysozyme.

The invention methods for producing multimeric proteins include immunoglobulins, such as antibodies, and antigen binding fragments thereof. Thus, in one embodiment of the present invention, the multimeric protein is an

immunoglobulin, wherein the first and second heterologous polypeptides are an immunoglobulin heavy and light chains respectively. Illustrative examples of this and other aspects of the present invention for the production of heterologous multimeric polypeptides in avian cells are fully disclosed in U.S. Patent Application No. 09/877,374, filed June 8, 2001, by *Rapp*, published as US-2002-0108132-Al on August 8, 2002, and U.S. Patent Application No. 10/251,364, filed September 18, 2002, by *Rapp*, both of which are incorporated herein by reference in their entirety.

Accordingly, the invention further provides immunoglobulin and other multimeric proteins that have been produced by transgenic avians of the invention.

In various embodiments, an immunoglobulin polypeptide encoded by the transcriptional unit of at least one expression vector may be an immunoglobulin heavy chain polypeptide comprising a variable region or a variant thereof, and may further comprise a D region, a J region, a C region, or a combination thereof. An immunoglobulin polypeptide encoded by an expression vector may also be an immunoglobulin light chain polypeptide comprising a variable region or a variant thereof, and may further comprise a J region and a C region. The present invention also contemplates multiple immunoglobulin regions that are derived from the same animal species, or a mixture of species including, but not only, human, mouse, rat, rabbit and chicken. In preferred embodiments, the antibodies are human or humanized.

In other embodiments, the immunoglobulin polypeptide encoded by at least one expression vector comprises an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region, and a linker peptide thereby forming a single-chain antibody capable of selectively binding an antigen.

Another aspect of the present invention provides a method for the production in an avian of an heterologous protein capable of forming an antibody suitable for selectively binding an antigen. This method comprises a step of producing a transgenic avian incorporating at least one transgene, wherein the transgene encodes at least one heterologous polypeptide selected from an immunoglobulin heavy chain variable region, an immunoglobulin heavy chain comprising a variable region and a constant region, an immunoglobulin light chain variable region, an immunoglobulin

light chain comprising a variable region and a constant region, and a single-chain antibody comprising two peptide-linked immunoglobulin variable regions.

In one embodiment of this method, the isolated heterologous protein is an antibody capable of selectively binding to an antigen which may be generated by combining at least one immunoglobulin heavy chain variable region and at least one immunoglobulin light chain variable region, preferably cross-linked by at least one disulfide bridge. The combination of the two variable regions generates a binding site that binds an antigen using methods for antibody reconstitution that are well known in the art.

The present invention also encompasses immunoglobulin heavy and light chains, or variants or derivatives thereof, to be expressed in separate transgenic avians, and thereafter isolated from separate media including serum or eggs, each isolate comprising one or more distinct species of immunoglobulin polypeptide. The method may further comprise the step of combining a plurality of isolated heterologous immunoglobulin polypeptides, thereby producing an antibody capable of selectively binding to an antigen. In this embodiment, for instance, two or more individual transgenic avians may be generated wherein one transgenic produces serum or eggs having an immunoglobulin heavy chain variable region, or a polypeptide comprising such, expressed therein. A second transgenic animal, having a second transgene, produces serum or eggs having an immunoglobulin light chain variable region, or a polypeptide comprising such, expressed therein. The polypeptides from two or more transgenic animalsmay be isolated from their respective sera and eggs and combined in vitro to generate a binding site capable of binding an antigen.

Examples of therapeutic antibodies that can be used in methods of the invention include but are not limited to HERCEPTIN® (Trastuzumab) (Genentech, CA) which is a humanized anti-HER2 monoclonal antibody for the treatment of patients with metastatic breast cancer; REOPRO® (abciximab) (Centocor) which is an anti-glycoprotein IIb/IIIa receptor on the platelets for the prevention of clot formation; ZENAPAX® (daclizumab) (Roche Pharmaceuticals, Switzerland) which is an immunosuppressive, humanized anti-CD25 monoclonal antibody for the

prevention of acute renal allograft rejection; PANOREX™ which is a murine anti-17-IA cell surface antigen IgG2a antibody (Glaxo Wellcome/Centocor); BEC2 which is a murine anti-idiotype (GD3 epitope) IgG antibody (ImClone System); IMC-C225 which is a chimeric anti-EGFR IgG antibody (ImClone System); VITAXIN™ which is a humanized anti-αVβ3 integrin antibody (Applied Molecular Evolution/MedImmune); Campath 1H/LDP-03 which is a humanized anti CD52 IgG1 antibody (Leukosite); Smart M195 which is a humanized anti-CD33 IgG antibody (Protein Design Lab/Kanebo); RITUXAN™ which is a chimeric anti-CD2O IgG1 antibody (IDEC Pharm/Genentech, Roche/Zettyaku); LYMPHOCIDE™ which is a humanized anti-CD22 IgG antibody (Immunomedics); ICM3 is a humanized anti-ICAM3 antibody (ICOS Pharm); IDEC-114 is a primatied anti-CD80 antibody (IDEC Pharm/Mitsubishi); ZEVALIN™ is a radiolabelled murine anti-CD20 antibody (IDEC/Schering AG); IDEC-131 is a humanized anti-CD40L antibody (IDEC/Eisai); IDEC-151 is a primatized anti-CD4 antibody (IDEC); IDEC-152 is a primatized anti-CD23 antibody (IDEC/Seikagaku); SMART anti-CD3 is a humanized anti-CD3 IgG (Protein Design Lab); 5G1.1 is a humanized anti-complement factor 5 (CS) antibody (Alexion Pharm); D2E7 is a humanized anti-TNF-α antibody (CATIBASF); CDP870 is a humanized anti-TNF-α Fab fragment (Celltech); IDEC-151 is a primatized anti-CD4 IgG1 antibody (IDEC Pharm/SmithKline Beecham); MDX-CD4 is a human anti-CD4 IgG antibody (Medarex/Eisai/Genmab); CDP571 is a humanized anti-TNF-α IgG4 antibody (Celltech); LDP-02 is a humanized anti-α4β7 antibody (LeukoSite/Genentech); OrthoClone OKT4A is a humanized anti-CD4 IgG antibody (Ortho Biotech); ANTOVA™ is a humanized anti-CD40L IgG antibody (Biogen); ANTEGREN™ is a humanized anti-VLA-4 IgG antibody (Elan); and CAT-152 is a human anti-TGF-β_2 antibody (Cambridge Ab Tech).

Production of Exogenous Protein by Transgenic Avians

Methods for the production of heterologous protein by the avian oviduct and the production of eggs which contain heterologous protein involve providing a suitable vector and introducing the vector into embryonic blastodermal cells so that the vector can integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells produced in the

previous steps. Deriving a mature transgenic avian from the blastodermal cells optionally involves transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. Another alternative is to transfer the transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In an alternative embodiment, the cells of a blastodermal embryo are transfected or transduced with the vector directly within the embryo. It is contemplated, for example, that the recombinant nucleic acid molecules of the present invention may also be introduced into a blastodermal embryo by direct microinjection of the DNA into a Stage X or earlier embryo that had been removed from the oviduct. The egg is then returned to the bird for shell development and laying. The resulting embryo is allowed to develop and the chick allowed to mature.

In either case, the transgenic bird so produced from the transgenic blastodermal cells is known as a "founder" Some founders can be chimeric or mosaic birds if, for example, microinjection does not deliver nucleic acid molecules to all of the blastodermal cells of an embryo. Some founders will carry the transgene in the tubular gland cells in the magnum of their oviducts and will express the exogenous protein encoded by the transgene in their oviducts. If the exogenous protein contains the appropriate signal sequences, it will be secreted into the lumen of the oviduct and onto the yolk of an egg.

Some founders are germ-line founders. A germ-line founder is a founder that carries the transgene in genetic material of its germ-line tissue, and may also carry the transgene in oviduct magnum tubular gland cells that express the exogenous protein. Therefore, in accordance with the invention, the transgenic bird will have tubular gland cells expressing the exogenous protein and the offspring of the transgenic bird will also have oviduct magnum tubular gland cells that express the exogenous protein. (Alternatively, the offspring express a phenotype determined by expression of the exogenous gene in a specific tissue of the avian.)

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals,

diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β-casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention. Other possible proteins to be produced include, but are not limited to, albumin, α-1 antitrypsin, antithrombin III, collagen, factors VIII, IX, X (and the like), fibrinogen, hyaluronic acid, insulin, lactoferrin, protein C, erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), tissue-type plasminogen activator (tPA), feed additive enzymes, somatotropin, and chymotrypsin. Immunoglobulins and genetically engineered antibodies, including immunotoxins which bind to surface antigens on human tumor cells and destroy them, can also be expressed for use as pharmaceuticals or diagnostics.

One aspect of the present invention, therefore, concerns transgenic birds, such as chickens, comprising a recombinant nucleic acid molecule of the present invention and which preferably (though optionally) express a heterologous gene in one or more cells in the animal. Suitable methods for the generation of transgenic avians having heterologous DNA incorporated therein are described, for example, in WO 99/19472 to Ivarie *et al.*; WO 00/11151 to Ivarie *et al.*; and WO 00/56932 to Harvey *et al.*, all of which are incorporated herein by reference in their entirety.

In various embodiments of the transgenic bird of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *trans*-acting factors acting on the ovalbumin transcriptional regulatory region of the present invention and which control gene expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences. The inclusion of a 5' MAR, and optionally a 3' MAR region, in the novel nucleic acid molecules of the present invention will allow the heterologous expression unit to escape all, or in part, the chromosomal positional

effect and therefore be expressed at a more uniform level in transgenic tissues that received the transgene by a route other than through germ line cells.

In various embodiments of the present invention the transgenic avians comprise a recombinant nucleic acid comprising SEQ ID NO: 1, a truncated variant of SEQ ID NO: 1, or the complement thereof.

In one embodiment of the present invention, the transgenic avian is selected from the group consisting of a chicken, a turkey, a duck, a goose, a quail, a pheasant, a ratite, an ornamental bird or a feral bird. In a preferred embodiment, the avian is a chicken.

In various embodiments, the transgenic avian produces the heterologous polypeptide in the serum or an egg white, or both.

The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patents and patents cited throughout the present application are hereby incorporated by reference in their entireties.

It will be apparent to those skilled in the art that various modifications, combinations, additions, deletions and variations can be made in the present invention without departing from the scope or spirit of the invention. For instance, features illustrated or described as part of one embodiment can be used in another embodiment to yield a still further embodiment. It is intended that the present invention covers such modifications, combinations, additions, deletions and variations as come within the scope of the appended claims and their equivalents.

Example 1:Construction of a complete Ovalbumin locus from two overlapping BACs.

A complete ovalbumin locus BAC was created from two overlapping BACs that together contained the complete ovalbumin locus, as shown in Fig. 6. The nucleotide sequences of BAC 120 and BAC 77 are in opposite directions with respect to the vector backbone pECBAC1.

BAC 120 was digested with Not I and a 145 kb fragment was re-cloned, but in the reversed orientation (flipped), into Not I digested vector backbone pECBAC1. This resulted in a deletion of a region of approximately 11.5 kb from the 5' end of the insert sequence of BAC 120 and which was upstream of the DNase I sensitivity region. The reversed BAC 120 'flip' and BAC 77 clones were digested with Srf I and RARE digested using an oligonucleotide targeted to an EcoRI site within ovalbumin. 5' and 3' fragments were isolated by CHEF gel electrophoresis, and ligated together to yield the complete contiguous ovalbumin genomic locus BAC.

Example 2: Expression of a heterologous gene by a chicken ovalbumin locus BAC

cDNA constructs encoding immunoglobulin light-chain and heavy-chains of a human IgG_1 kappa monoclonal antibody were inserted in-frame with the ovalbumin translation start site of separate ovalbumin locus-containing BACs, as shown in Fig 3. The immunoglobulin chain-encoding cDNAs were first inserted into a plasmid that contained a 2.7 kb EcoRI fragment from the ovalbumin gene and which included the ovalbumin start site. The resulting vector was then digested with restriction endonuclease EcoR1 and cloned into an approximately 195 kb ovalbumin BAC which had been subjected to EcoR1 recA-assisted restriction endonuclease (RARE) digestion as described by Boren et al., 1996, *Prot. Sci.* 5,: 2479-2484 and incorporated herein by reference in its entirety.

Transgenic birds were created by cytoplasmic co-microinjection of human light-chain and heavy chain BACs (figure b) followed by ovum transfer as described in U.S Patent Application Serial No.10/251,364 incorporated herein by reference in its entirety.

A hen carrying these constructs was grown to sexual maturity. Eggs were collected and the egg white material was assayed for the expressed human monoclonal antibody using sandwich ELISA as described by *Harlow et al., Antibodies: a Laboratory Manual. 1988, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory. Xiii* incorporated herein by reference in its entirety. The human monoclonal antibody was captured by a goat anti-human kappa chain specific

monoclonal antibody and quantified with an alkaline phosphatase conjugated goat anti-human gamma detection antibody. Hen # AA698 expressed up to 1025 pg of human monoclonal antibody per ml of egg white.

Example 3: Expression of a heterologous gene by a chicken ovalbumin locus BAC.

The open reading frame of the firefly luciferase gene was inserted into the ovalbumin translation start site of an ovalbumin locus BAC as shown in Fig 3. The luciferase gene was inserted into a plasmid that contained a 2.7 kb EcoR1 fragment from the ovalbumin gene and which includes the ovalbumin start site. The resulting vector was then digested with EcoRI and cloned into an approximately 195kb ovalbumin BAC which had been subjected to EcoR1 recA-assisted restriction endonuclease (RARE) digestion as described by Boren et al., 1996, *Prot. Sci.* 5,: 2479-2484 and incorporated herein by reference in its entirety.

Primary tubular gland cells isolated from the oviduct of laying quail (Sanders and McKnight, Endocrinology 116, 398-405(1985)), were transfected using the ovalbumin-luciferase construct or with a negative control CMV-IFN construct. Luciferase activity in cell extracts was analyzed two days post transfection (Table 1).

Table1

DNA	RLU
CMV-IFN	60
Ovalbumin Luciferase	274

Example 4: Basic Local Alignment Search Tool (BLAST) Analysis of the Complete Ovalbumin Promoter Sequence (SEQ ID NO: 1)

The complete approximately 195kb ovalbumin promoter sequence (SEQ ID NO: 1) was submitted to the National Center for Biotechnology Information for BLAST alignments with database sequences. Further analysis was by using the GenScan and MARWIZ software. Percent identities between the ovalbumin gene region sequence (SEQ ID NO: 1) and corresponding known ovalbumin promoter features are listed in Table 2 below.

Table 2:

Nucleotide positions of identifiable elements in the region of the chicken genomic within BACs 120, 77 and 26

Nucleotide Positions[a]	Domain Identity
5963-1	**Q**[b]
9730-9922	CR1
10772-11935	CpG Island
18914-19088	CR1-GG
20106-20921	CR1-GG
39975-24820	**R ATPase**
41119-41177	CR1-GG
41586-41700	CR1-GG
41701-41800	MAR element
42221-42742	CpG Island
43505-46990	**S Gene**
50017-51427	**T Gene**
56001-56201	MAR-like element
56501-56901	MAR-like element
64599-71919	**U Serpin Gene**
58401-58701	MAR-like element
74883-75634	CR1-GG
75420-75634	CR1b
76251-76451	MAR-like element
80151-80451	MAR-like element
81125-94938	**V Serpin Gene**
81832-82120	CR1
85473-85922	CR1-GG
88654-88797	CRI-GG
90120-90167	CRI-GG
96401-96800	MAR element
97884-97965	Y:OV-1 element

99080-99107	SDRE element
100602-107839	**X Gene**
110247-111200	CR1-GG
114779-121099	**Y Gene**
117849-118132	CR1-GG
131729-139290	**Ovalbumin**
144651-144850	MAR element
147721-155849-	**W Gene**
150801– 151600	MAR element
156581-157181	MAR-like element
157081-157581	MAR-like element
157132-157331	MAR-like element
159095-165114	**MENT**
163701–164100	MAR element
171633-180432	**Z1**
183204-190418	**Z2**
186201–186590	MAR element
190101–190800	MAR element
192078-195101	**Z3**

[a]Nucleotide positions of protein encoding regions are from the beginning of the first exon to the end of the poyadenylation signal-exons are shown in Fig. 1

[b]Protein coding regions are given in bold

Although preferred embodiments of the invention have been described using specific terms, devices, and methods, such description is for illustrative purposes only. The words used are words of description rather than of limitation. It is to be understood that changes and variations may be made by those of ordinary skill in the art without departing from the spirit or the scope of the present invention, which is set forth in the following claims. In addition, it should be understood that aspects of the various embodiments may be interchanged both in whole or in part.